Proceedings of the International Conference , "Computational Systems for Health & Sustainability"
17-18, April, 2015 - by R.V.College of Engineering,
Bangalore,Karnataka,PIN-560059,INDIA

People Counting System- A Review

Kowcika.A
Asst.Professor, DEPARTMENT OF CSE
R.V.College of Engineering, BANGALORE 560059 India

Abstract: The method of people counting has been developed using two different approaches. In the direct approach (also called detection-based), people in the scene are first individually detected, using some form of segmentation and object detection, and then counted. In the indirect approach (also called map-based), instead, counting is performed using the measurement of some feature that does not require the separate detection of each person in the scene. This paper focuses on human detection first and comprehensive review on the two methods of people counting.

INTRODUCTION

People count Extraction is a crucial and challenging problem in visual surveillance. For instance an accurate and real-time estimation of number of people present in a shopping mall provides significant customer information for managers. Further, automatic monitoring of the people counts in public areas is essential for safety control and urban planning. In recent years, though this problem domain is well explored; yet the following issues are to be resolved such as videos that consist of stationary objects, complicated background, and low resolution. However, real scenes always include both moving/stationary human beings, the complicated background, and most videos in a visual surveillance system have a relatively low resolution. The method of People count extraction falls under two categories based on the Purpose [1]. (a).Estimate the number of people in a region of interest (ROI counting). (b).To count the crowds across a line of interest (LOI counting).

The two different methods for extracting the number People in crowds are Detection-based methods and Map-based methods. Detection based methods determine the number of people and their locations using some form of segmentation and human detection such as head, shoulder, head-tops .Map-based methods exploits the relationship between the people and local/global features from the image. [2].

Some of the applications to show the importance of people count are as follows,

- Occupancy limit in a building.
- To actively manage city services.
- Allocate resources for public events.
- To aid with crowd control during rallies.
- Crowd disaster prevention.
- Group behavior modeling.
- To provide usage statistics to metro operators.
- An accurate and real-time estimation of people in a shopping mall can provide valuable information for managers.
- Tourists flow estimation.

Challenges of people counting:

Crowd scenes remain particularly challenging for the detection and tracking tasks due to the following reasons,

a) Heavy occlusions, high person densities and significant variation in people's appearance.

b) The ambiguous appearance of body parts, e.g. the head of one person could be similar to a shoulder of a near-by person.

c) People must be moving, the background must be simple, and the image resolution must be high.

d) Varying background and lighting condition.

e) Inter-object occlusion, Scene occlusion , Cluttered background

f) People can have a wide range of scales

Inter-object occlusion

Scene occlusion

Cluttered background

Uneven illumination

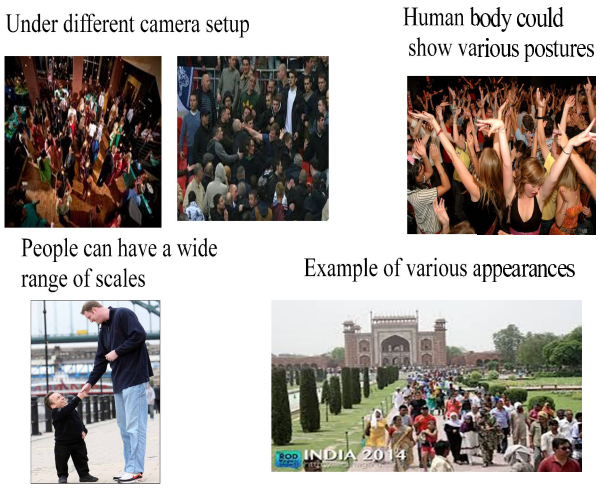

Figure.1

Human Detection System:

Human Detection can be performed as two step process. The first step is to detect the moving object followed by classification. Object detection could be performed using background subtraction, optical flow and spatio-temporal filtering techniques. A moving object could be classified as a human being using shape-based, texture-based or motion-based features. A comprehensive review of available techniques for detecting human beings in surveillance videos is presented in this paper [3].

The people count can be achieved by implementing following modules,

- Background Subtraction
- Physical/High Feature Extraction or detection of body parts (Detection based or Map based method)
- Applying Supervised/Unsupervised Learning Algorithms

Detection Based Method for People count:

In the existing approaches extraction of people count is implemented using the algorithms such as EM. For example Hou et al used the Expectation Maximization (EM)-based algorithms to locate individuals in a low resolution scene. Though this approach is accurate in both counting and detection it does not distinguish human and non-human objects more accurately in a low resolution videos [2].

Rodriguez et al explored constraints imposed by the crowd density and formulate person detection as the optimization of a joint energy function combining crowd density estimation and the localization of individual people. They demonstrated how the optimization of such an energy function significantly improves person detection and tracking in crowds. Currently video frames are processed individually and obtained detections are tracked in post-processing that is during tracking but not in detecting [4].

Xu et al presents a counting system which consists of four modules: foreground extraction, head-shoulder component detection, tracking and trajectory analysis. It reduced the computation costs and cope with various complex surveillance situations for foreground extraction, an adaptive components number selection strategy for mixture of Gaussians model is proposed. Pedestrians are detected by their head shoulders, because this part is less varied and less likely occluded from a downward-slope view. Each pedestrian is tracked through consecutive frames using the Kalman filter techniques and cost function. The resulting trajectories are analyzed to count people entering or leaving the scene But the limitations of this approach lies in handling Occlusions [7].

Merad et al detection based method where the skeleton silhouette is computed and decomposed into a set of segments corresponding to the head, torso and limbs. This structure captures the minimal information about the skeleton shape. No assumption is made about the view point; this is done after the head pose process. Several results present the efficiency of the labeling process, particularly its structural properties for the detection of heads within a crowd. The experimental results proved that the challenges (e) and (f) are not faced by this approach [8].

Bhondi et al proposed a system for automatic people counting in crowded environments. The proposed approach is a counting-by-detection method based on depth imagery. It is designed to be deployed as an autonomous appliance for crowd analysis in video surveillance application scenarios. This system performs foreground/background segmentation on depth image streams in order to coarsely segment persons, and then depth information is used to localize head candidates which are then tracked in time on an automatically estimated ground plane. An extensive comparative evaluation is discussed. Though this method handled crowds, queuing and groups well the challenges (b), (c) and (d) are not highlighted [9].

Map Based Method for People count:

Zhang et al used a complex network-based algorithm to detect interest points and extract the global texture features in scenarios. After this point degree matrices is computed and statistical measures are obtained. Then it establishes a mapping between the moving interest point's features and the number of people in a crowd scene. This method is much effective only for specific test cases and also towards Occlusion [5].

Hou et al proposed a method in which is more accurate in both counting and detection. First, post-processing steps are performed on background subtraction results to estimate the number of people in a complicated scene, which includes people who are moving only slightly. Second, an Expectation

Maximization (EM)-based method has been developed to locate individuals in a low resolution scene. In this process, a new cluster model is used to represent each person in the scene. This process does not require a very accurate foreground contour. Third, the number of people is used as a priori for locating individuals based on feature points. Hence, the methods for estimating the number of people and for locating individuals are connected. However, this approach distinguished human and non-human objects more accurately; a high-resolution video is needed to provide sufficient data [6].

Hussain et al [10] employ pixel-based approaches for feature extraction and back- propagation neural network for people counting and estimation for crowd density at Masjidal-Haram (Mosque). The experimental results indicate high accuracy and performance level for very low to moderate crowd. However, when the density of the crowd increases, the detection rate tends to decrease due to occlusion level changes.

CONCLUSION

The purpose of this review article is to present two different approaches, the direct approach (also called detection-based), people in the scene are first individually detected, using some form of segmentation and object detection, and then counted. In the indirect approach (also called map-based), instead, counting is performed using the measurement of some feature that does not require the separate detection of each person in the scene. This paper focuses on human detection first and comprehensive review on the two methods of people counting. The idea of presenting this review is the first step to do more research in this direction to come out with some innovative research findings soon.

ACKNOWLEDGMENT

The author is grateful to Research Guide, Dr.S.Sridhar, Professor and Dean-Cognitive and Central Computing, of our college for the guidance to prepare this review article and also thankful to Dr.G.Shobha, HoD/CSE and Dr.N.K.Srinath, PG Dean (CSE) for the support offered.

REFERENCES

[1]. Li, Jingwen, L. Huang, C.Liu,"Robust people counting in video surveillance: Dataset and system." in 8th IEEE Int. Conf. on Advanced Video and Signal-Based Surveillance, pp. 54-59, 2011.

[2]. Hou, Ya-Li, Grantham KH Pang. "People counting and human detection in a challenging situation", in Part A: Systems and Humans, IEEE Trans. on Systems, Man and Cybernetics, 41, no. 1, 24-33, 2011.

[3]. Paul, M., Haque, S.M.E., Chakraborty, S,"Human detection in surveillance video and its applications — a review" EURASIP Journal on Advanced in Signal Processing 2013, pp.1–16 ,2013.

[4]. Rodriguez, Mikel, Ivan Laptev, Josef Sivic, and J-Y. Audibert. "Density-aware person detection and tracking in crowds." In Computer Vision (ICCV), 2011 IEEE International Conference on, pp. 2423-2430. IEEE, 2011.

[5]. Zhang, Yueguo, L.Dong, Jianhua Li, Shenghong Li, and Z. J. Gao. "A complex network-based approach to estimating the number of people in video surveillance", in IEEE Int. Symp. on Broadband Multimedia Systems and Broadcasting, pp. 1-4, 2013.

[6]. Hou, Ya-Li, Grantham KH Pang. "People counting and human detection in a challenging situation", in Part A: Systems and Humans, IEEE Trans. on Systems, Man and Cybernetics, 41, no. 1, 24-33, 2011.

[7]. Xu, Huazhong, Pei Lv, and Lei Meng. "A people counting system based on head-shoulder detection and tracking in surveillance video." Computer Design and Applications (ICCDA), 2010 International Conference on. Vol. 1. IEEE, 2010.

[8]. Merad, Djamel, K-E. Aziz, and Nicolas Thome. "Fast people counting using head detection from skeleton graph." Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on. IEEE, 2010.

[9]. Bondi, Enrico, et al. "Real-time people counting from depth imagery of crowded environments." Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on. IEEE, 2014.

[10]. Hussain et al,"CDES: A pixel-based crowd density estimation system for Masjid al-Haram." in Safety science 49, no. 6 (2011): 824-833.

